



Lessons Learned Bringing Up Early Adopter PCIe® 4.0 Links

Richard Solomon
Synopsys

Disclaimer



Presentation Disclaimer: All opinions, judgments, recommendations, etc. that are presented herein are the opinions of the presenter of the material and do not necessarily reflect the opinions of the PCI-SIG®.

Acknowledgements



- **Many thanks to Torrey Lewis for actually DOING the integration testing and providing screen shots and other data for the presentation!**

Agenda



- **Overview**
- **Testing Setup (Hardware/Software)**
- **Interesting PCIe 4.0[®] Features**
- **Failure Manifestations**
- **Resolution**

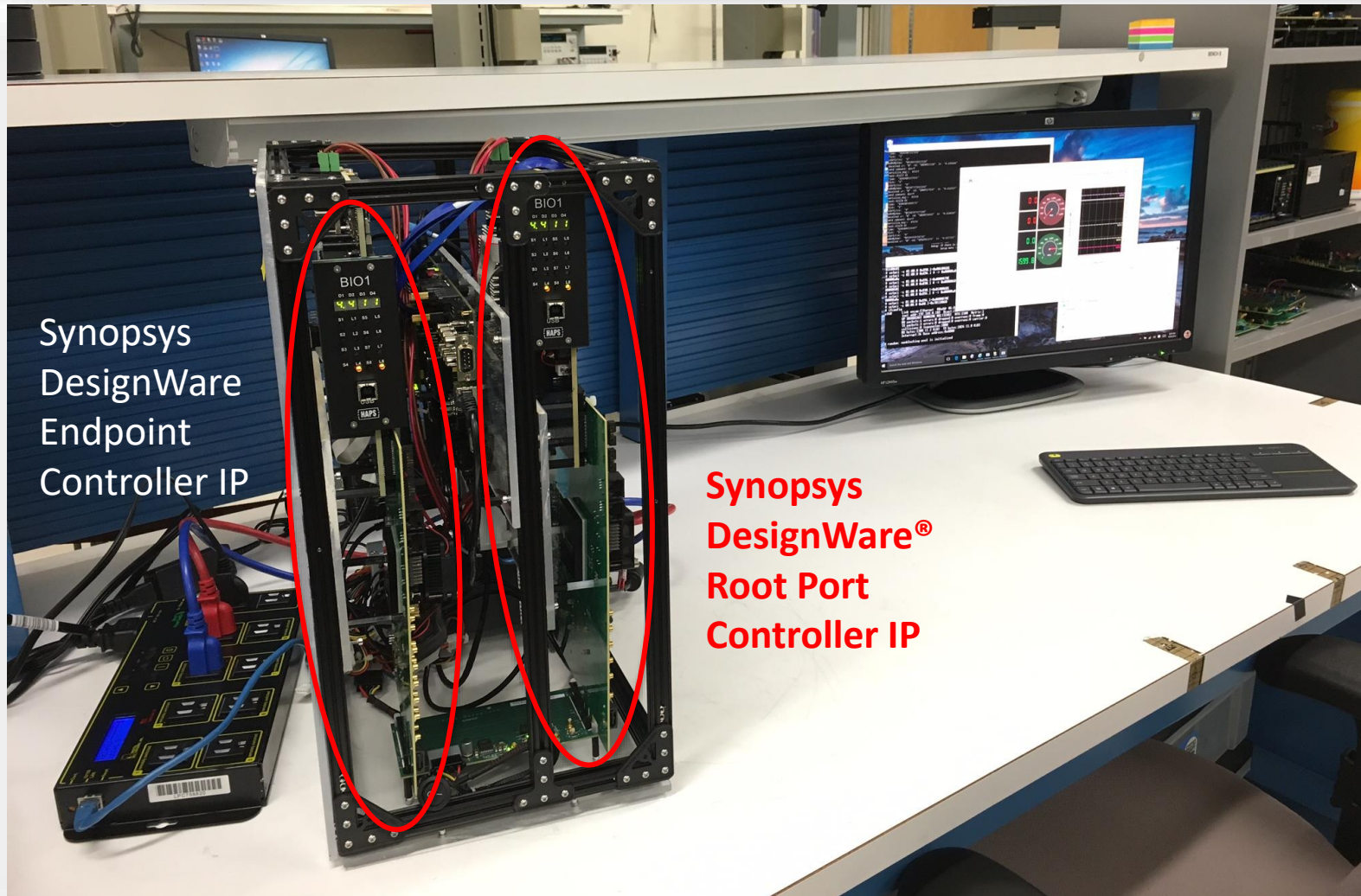
Testing Setup

Testing Setup (Hardware/Software)

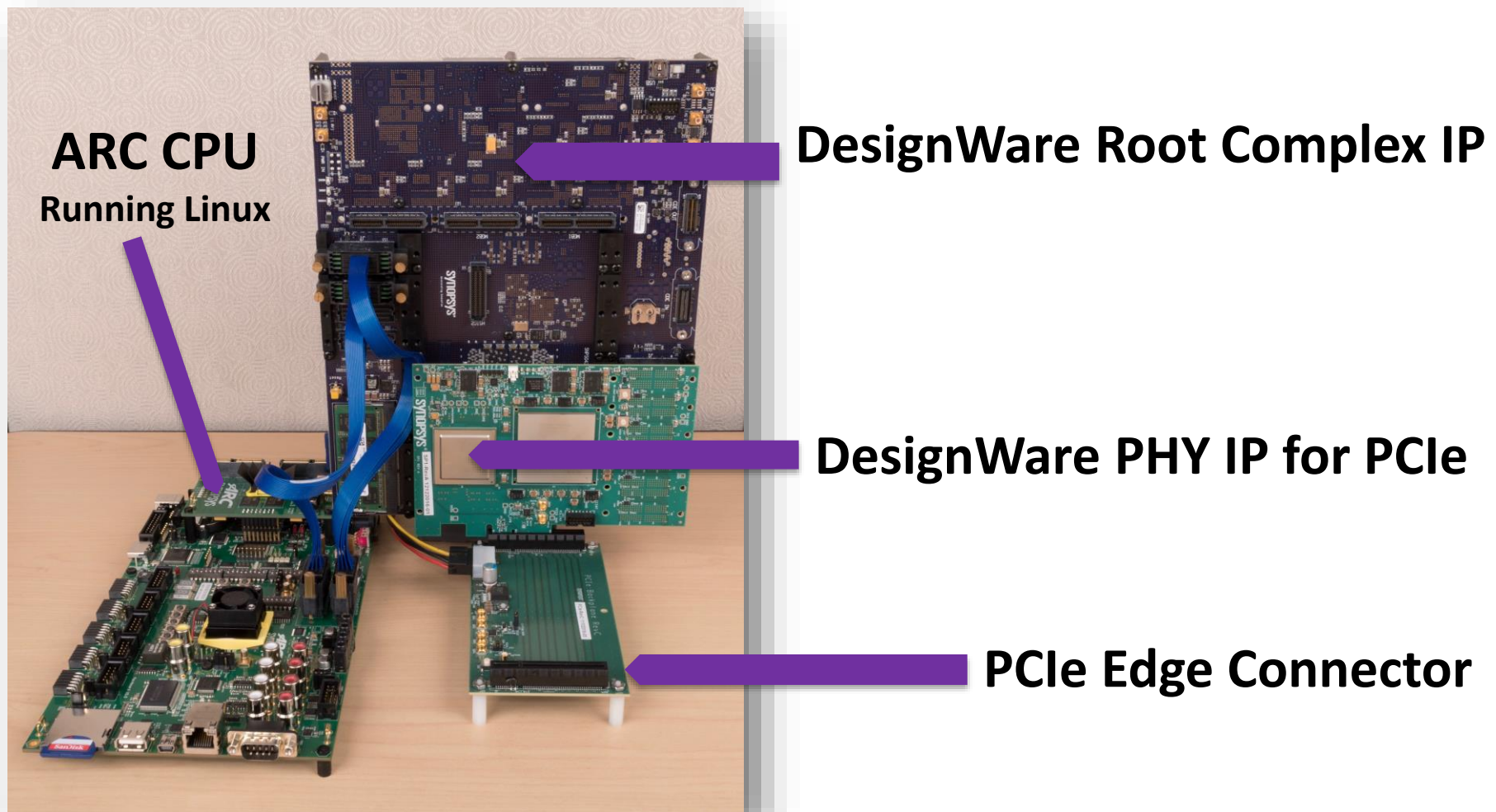


- **Synopsys PCIe 4.0 Root Complex**
 - Linux Running on ARC CPU single-board computer
 - Includes “RAS-DES” features
- **Synopsys PCIe 4.0 Endpoint**
 - Dedicated demo design with DMA engines, data movers, comparators, etc. to stimulate and respond to PCIe traffic
 - Includes “RAS-DES” features
- **Both on Synopsys HAPS Prototyping Platform**
 - FPGA-based
 - Includes Generic Embedded Logic Analyzer

Test Setup – Overview

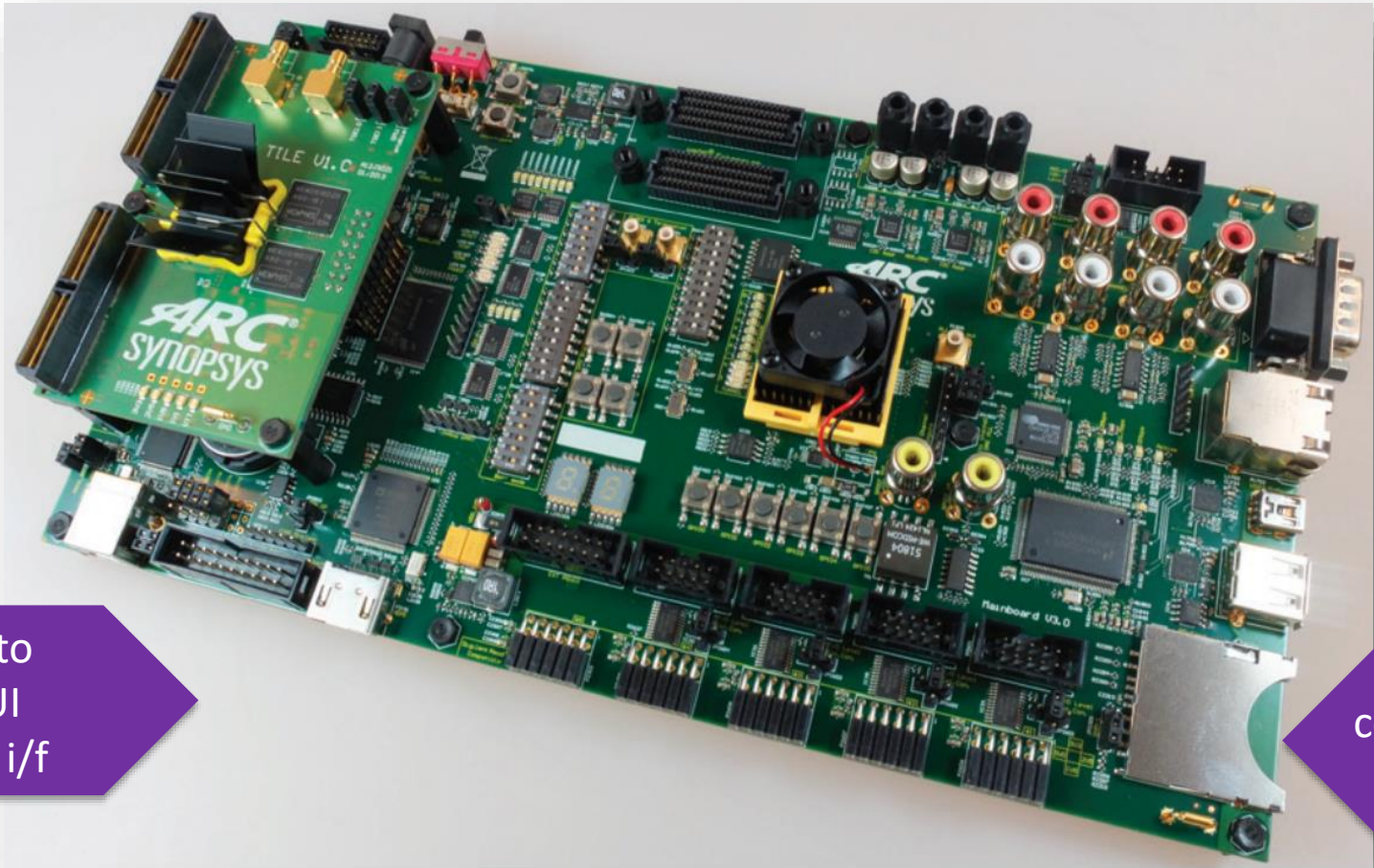


Test Setup – RC Side



Test Setup – CPU Details

ARC Software Development Platform

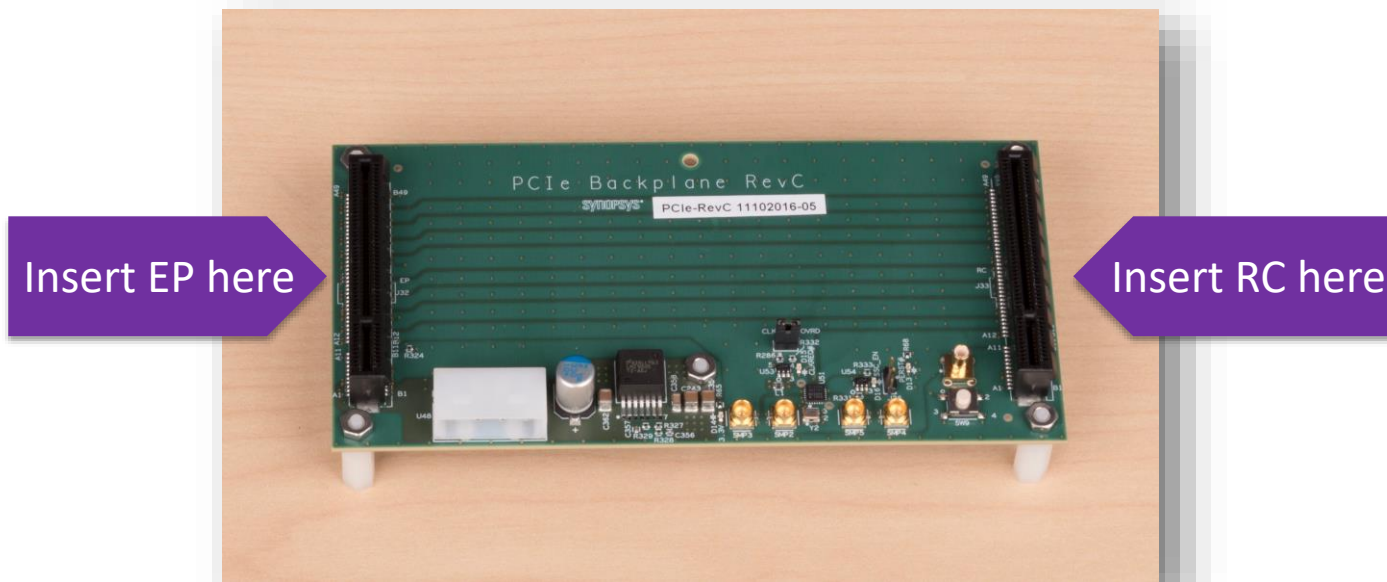


Connect to
PC for GUI
and shell i/f

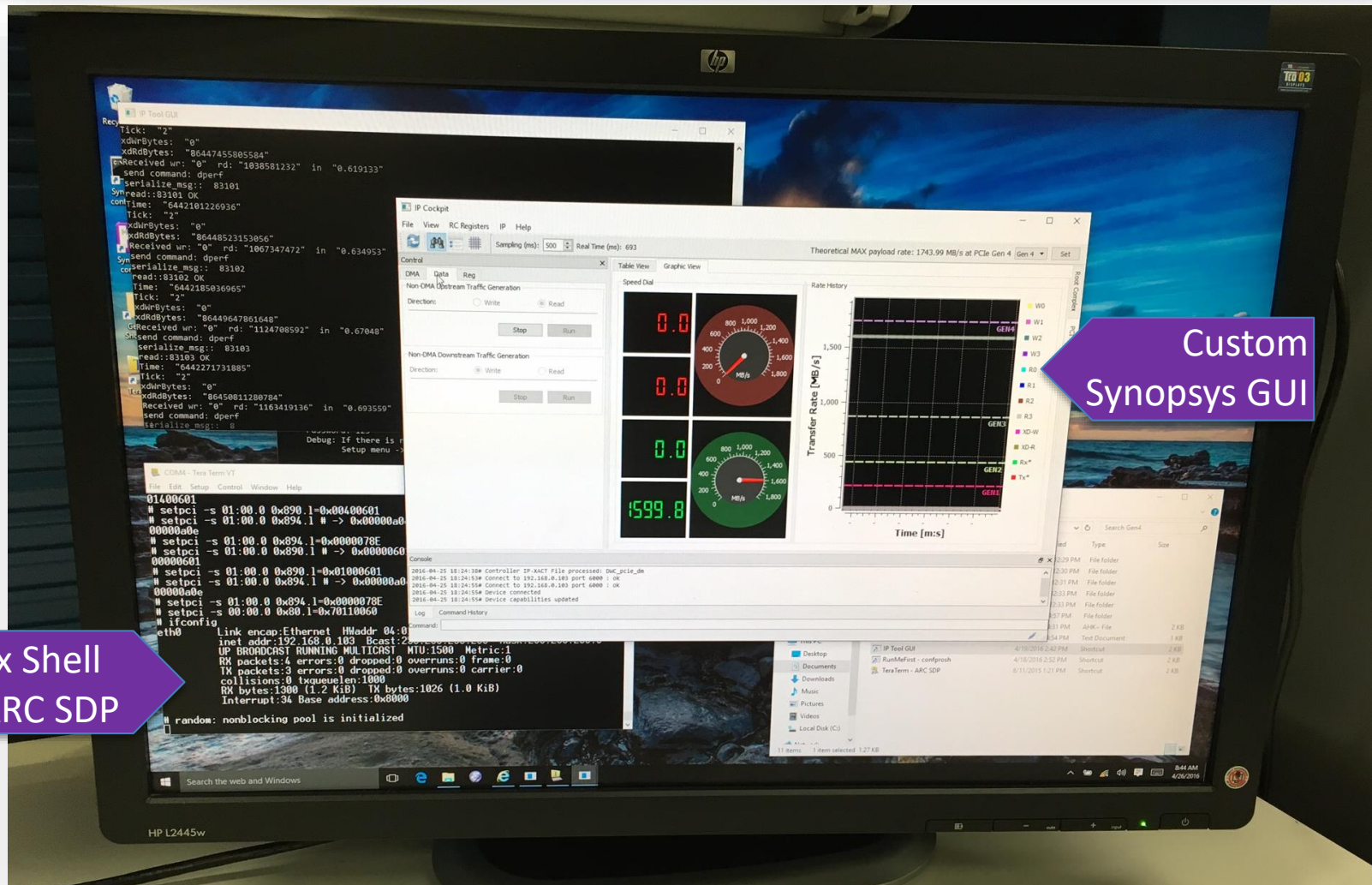
SD card
containing
Linux
ulmage

Test Setup – Backplane

Backplane connects RC IP HW to EP IP HW



Setup for Compliance Testing Software



Testing Setup (Hardware/Software Tools)



- **PCIe 4.0 Lane Margining**
 - Support still very early even in early adopter devices
 - Device-independent eye tool
- **External protocol analyzer**
 - Interposer card reduces signal integrity
- **On Chip logic analyzer**
 - In prototyping hardware
- **Design for Reliability (RAS-DES)**
 - Full details in Richard's 2015 PCI-SIG® Presentation
<https://members.pcisig.com/wg/PCI-SIG/document/8673?downloadRevision=active>

Lessons From the Past...



Tracking Link Reliability: Suggested Error Counters



- PHY Layer
 - ✓ Elasticity buffer overflow/underrun
 - ✓ Symbol Errors: disparity, decoding, loss of symbol lock
 - ✓ Receiver Error
 - ✓ N_FTS Timeout
 - ✓ Framing Error
 - ✓ Deskew Error
- Data Link Layer
 - ✓ Bad packet / LCRC
 - ✓ Replays: rollovers, timeouts
 - ✓ Rx/Tx Nak DLLP
- Transaction layer
 - ✓ FC Timeout
 - ✓ Poisoned TLP
 - ✓ ECRC Error
 - ✓ UR / CA
 - ✓ Completion Timeout

Lessons From the Past...



Tracking Link Reliability: Suggested Event Counters



- PHY Layer
 - ✓ Elasticity Buffer SKP Add/Delete
 - ✓ Speed change/Link width change
 - ✓ Entry to RECOVERY state
 - ✓ Ln state changes
- Data Link Layer
 - ✓ Tx/Rx Ack DLLP
 - ✓ Tx/Rx Update FC DLLP
 - ✓ Tx/Rx Nullified TLP
 - ✓ Rx Duplicate TLP
- Transaction layer
 - ✓ TX/RX packet types
 - ✓ FC credit exhaustion

Lessons From the Past...



Tracking Link Reliability: Suggested Status Information



- PHY Layer
 - ✓ SERDES interface signals
 - ✓ Lane detection
 - ✓ Polarity
 - ✓ SERDES power state
 - ✓ Deskew pointers
 - ✓ LTSSM states & registers
- Data Link Layer
 - ✓ Data Link Layer state
 - ✓ FC state
 - ✓ Sequence #s
 - ✓ Replay buffer depth
- Transaction layer
 - ✓ FC credit limits/usage
 - ✓ Packet queue depth

“Interesting” PCIe 4.0 Changes from 0.5 to 0.7

EIEOS Redefined for 16GT/s

- **Electrical Idle Exit Ordered Set**
- **0.5 was same as 8GT/s (eight 0s, eight 1s)**
- **0.7 changed to same frequency (sixteen each)**

Table 4-10: Electrical Idle Exit Ordered Set (EIEOS) for 8.0 GT/s and Above Data Rates

Symbol Numbers	Value	Description
0, 2, 4, 6, 8, 10, 12, 14	00h	Symbol 0: EIEOS Identifier A low frequency pattern that alternates between eight 0s and eight 1s.
1, 3, 5, 7, 9, 11, 13, 15	FFh	A low frequency pattern that alternates between eight 0s and eight 1s.

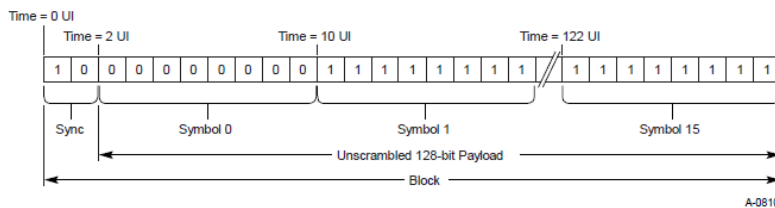
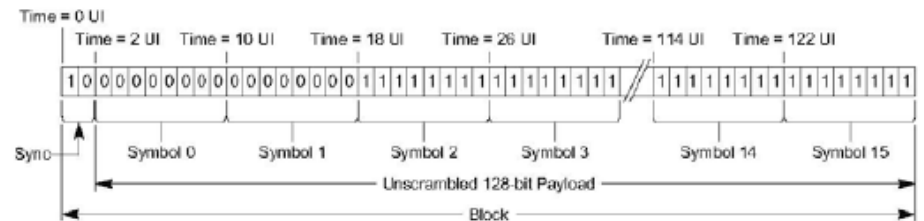


Figure 4-22: Electrical Idle Exit Ordered Set for 8.0 GT/s and Above Data Rates

Table 4-11: Electrical Idle Exit Ordered Set (EIEOS) for 16.0 GT/s Data Rate

Symbol Numbers	Value	Description
0, 1, 4, 5, 8, 9, 12, 13	00h	Symbol 0: EIEOS Identifier A low frequency pattern that alternates between sixteen 0s and sixteen 1s.
2, 3, 6, 7, 10, 11, 14, 15	FFh	A low frequency pattern that alternates between sixteen 0s and sixteen 1s.



(Electrical Idle Exit Ordered Set at 16.0 GT/s Data Rate)

A-0810B

- **Result: Mixed draft devices time out at 16GT/s due to never getting bit lock**

SKP Extended for 16GT/s



- **8GT/s SKP renamed to “Standard SKP” in 0.7**
- **0.7 adds “Control SKP” for Lane Margining**

Table 4-15: SKP Ordered Set with 128b/130b Encoding

Symbol Number	Value	Description
0 through (4*N - 1) [N can be 1 through 5]	AAh	SKP Symbol. Symbol 0 is the SKP Ordered Set identifier.
4*N	E1h	SKP_END Symbol. Signifies the end of the SKP Ordered Set after three more Symbols.
4*N + 1	00-FFh	(i) If LTSSM state is Polling.Compliance: AAh (ii) Else if prior block was a Data Block: Bit[7] = Data Parity Bit[6:0] = LFSR[22:16] (iii) Else: Bit[7] = ~LFSR[22] Bit[6:0] = LFSR[22:16]
4*N + 2	00-FFh	(i) If the LTSSM state is Polling.Compliance: Error_Status[7:0] (ii) Else LFSR[15:8]
4*N + 3	00-FFh	(i) If the LTSSM state is Polling.Compliance: ~Error_Status[7:0] (ii) Else: LFSR[7:0]

Table 4-17: Control SKP Ordered Set with 128b/130b Encoding

Symbol Number	Value	Description
0 through (4*N - 1) [N can be 1 through 5]	AAh	SKP Symbol. Symbol 0 is the SKP Ordered Set identifier.
4*N	78h	SKP_END_CTL Symbol. Signifies the end of the Control SKP Ordered Set after three more Symbols.
4*N + 1	00-FFh	Bit 7: Data Parity Bit 6: First Retimer Data Parity Bit 5: Second Retimer Parity Bits [4:0]: Margin CRC [4:0]
4*N + 2	00-FFh	Bit 7: Margin Parity Bits [6:0]: Refer to Section 4.2.13.1
4*N + 3	00-FFh	Bits [7:0]: Refer to Section 4.2.13.1

- **Result: Mixed draft devices slow due to unnecessarily looping through Recovery**

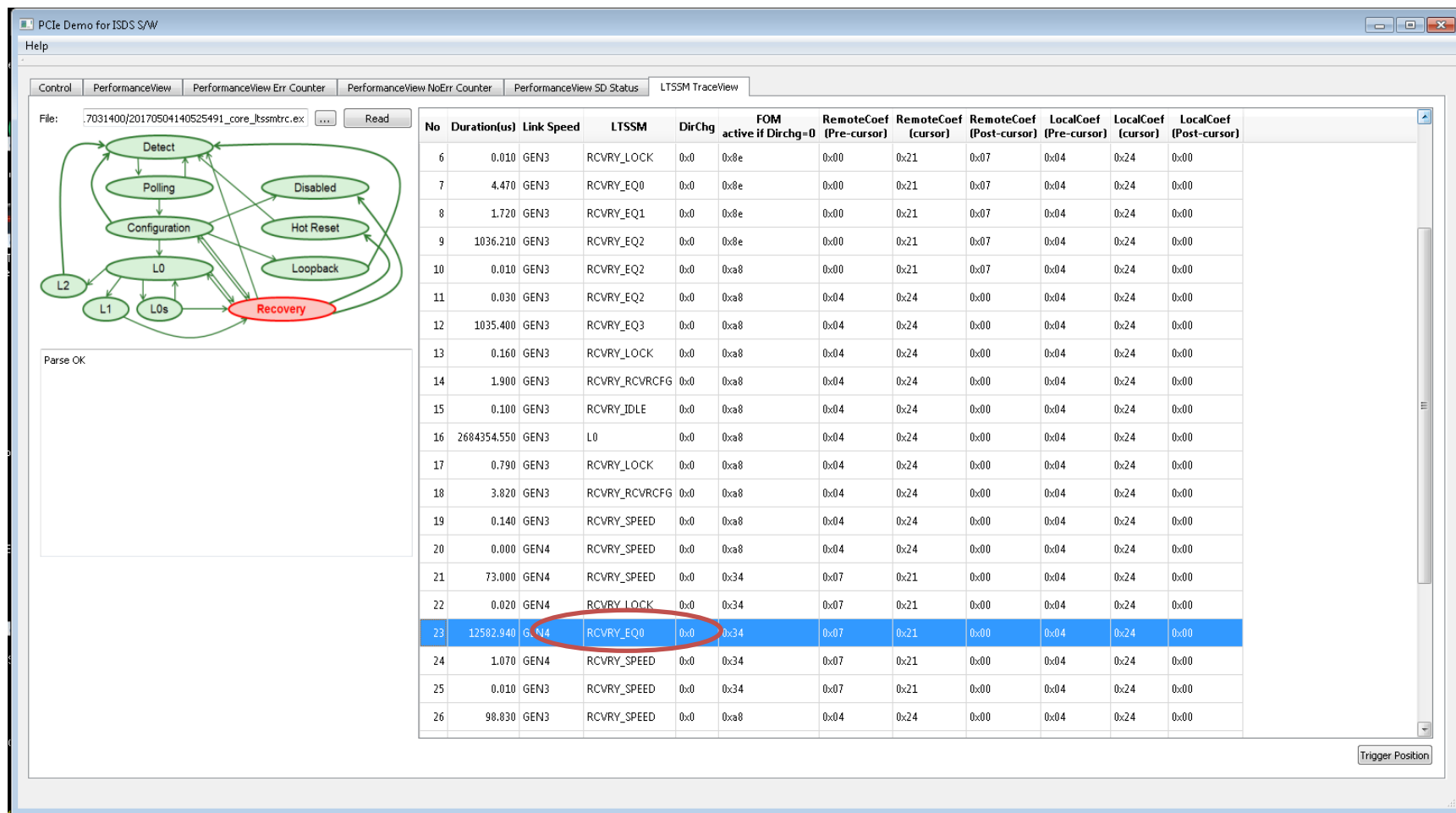
Lane Margining at the Receiver



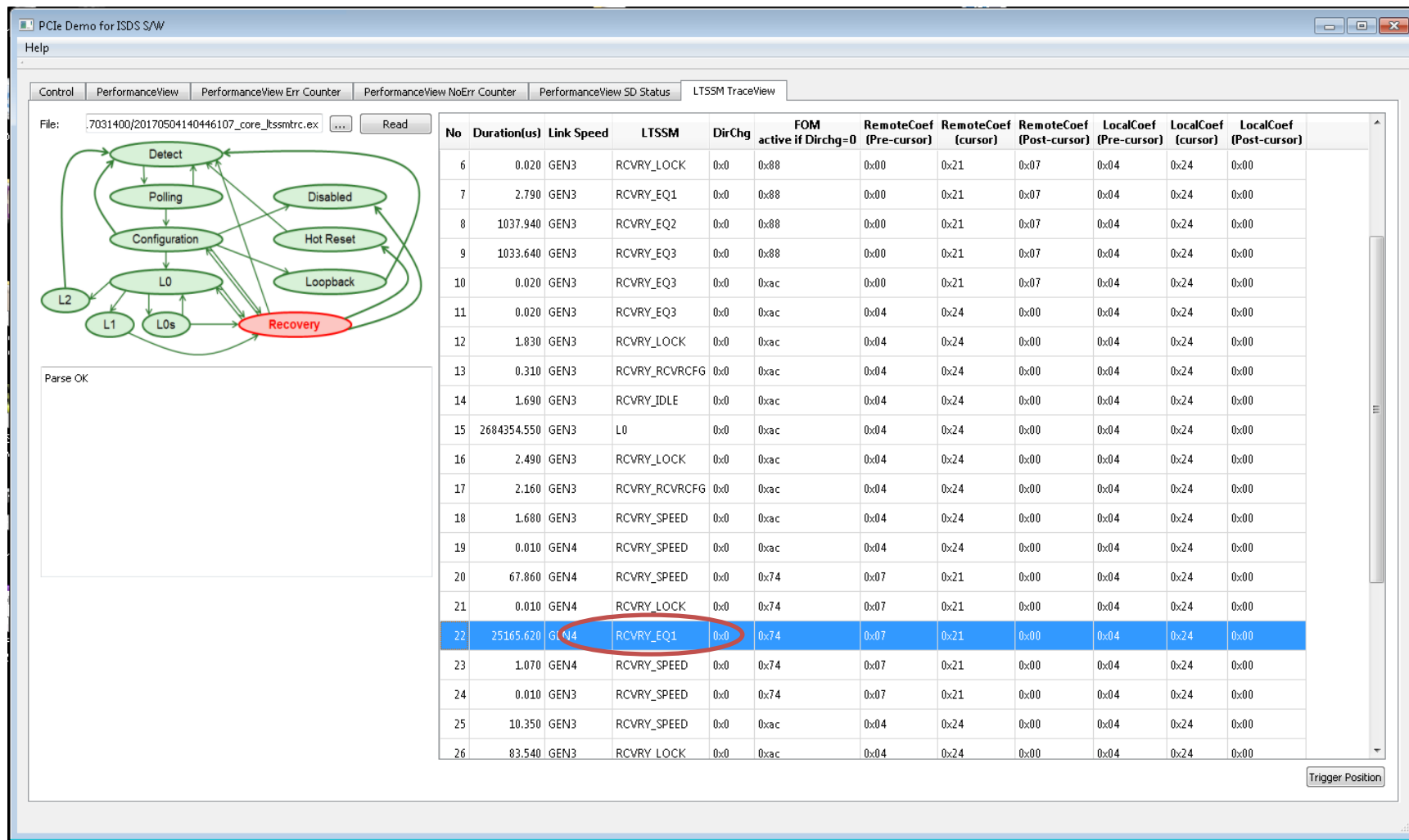
- **Allows the system to determine how close to “the edge” each lane is operating under real conditions**
- **Relies on information from receiving PHYs – conceptual “stepping” of sampling point in eye**
- **Will be controlled by software running on the host CPU on the motherboard**
 - NOT hardware autonomous, rudimentary “protocol” spec. for s/w
 - New capability structure in Root Complex/End Point/Switch
 - Retimer control via SKP ordered sets from downstream ports
- **Result is an approximation of the eye width and optionally the eye height for each lane**

Mismatch on Implemented Specification Revisions (0.5 vs 0.7)

0.5 EP / 0.7 RC (EP Side)



0.5 EP / 0.7 RC (RC Side)



Signal Integrity Issues

Unclean Link (RC Side)



PCIe Demo for ISDS S/W

Help

Control PerformanceView PerformanceView Err Counter PerformanceView NoErr Counter PerformanceView SD Status LTSSM TraceView

No Error Counter Per Lane(Max Value 15)

Lane0	Lane1	Lane2	Lane3	Lane4	Lane5	Lane6
EBUF SKP Add: 15	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0
EBUF SKIP DELL: 15	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0

LTSSM Counter(Max Value 4294967295)

L0 to Recovery Cnt: 101
L1 to Recovery Cnt: 0
Tx L0s: 0
Rx L0s: 0
ASPM L1 Reject: 0
L1 : 0
L1 CPM: 0
L1.1: 0
L1.2: 0
L1 short duration: 0
L1.2 abort: 0
L2: 0
Speed Change: 2
Link width Change: 0

Rx Counter(Max Value 4294967295)

Rx Memory Write: 0	Rx Ack DLLP: 0
Rx Memory Read: 0	Rx Update FC DLLP: 36556194
Rx Config Write: 0	Rx Nullified TLP: 0
Rx Config Read: 0	Rx Duplicate TLP: 0
Rx IO Write: 0	
Rx IO Read: 0	
Rx Completion wo data: 0	
Rx Completion w data: 0	
Rx Message TLP: 0	
Rx Atomic: 0	
Rx TLP with Prefix: 0	

Tx Counter(Max Value 4294967295)

Tx Memory Write: 0
Tx Memory Read: 0
Tx Config Write: 0
Tx Config Read: 0
Tx IO Write: 0
Tx IO Read: 0
Tx Completion wo data: 0
Tx Completion w data: 0
Tx Message TLP: 0
Tx Atomic: 0
Tx TLP with Prefix: 0

Unclean Link (RC Side)



PCIe Demo for ISDS S/W

Help

Control PerformanceView PerformanceView Err Counter PerformanceView NoErr Counter PerformanceView SD Status LTSSM TraceView

Common Error counter(Max Value 65535)

Detect EI Infer:	0
Receiver Error:	110
Rx Recovery Request:	0
N_FTS Timeout:	0
Framing Error:	106
Deskew Error:	105
BAD TLP:	0
LCRC Error:	0
BAD DLLP:	0
Replay Num Rollover:	0
Replay Timeout:	0
Rx Nak DLLP:	0
Tx Nak DLLP:	0
Retry TLP:	0
FC Timeout:	0
Poisoned TLP:	0
ECRC Error:	0
Unsupported Request:	0
Completer Abort:	0
Completion Timeout:	0

Error counter per Lane(Max Value 15)

Lane0	Lane1	Lane2	Lane3
EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0
EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0
Decode Error: 12	Decode Error: 0	Decode Error: 0	Decode Error: 0
Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0
SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0
SYNC Header Error: 5	SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0
Rx Valid de-assertion: 15	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0

Lane4	Lane5	Lane6	Lane7
EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0
EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0
Decode Error: 0	Decode Error: 0	Decode Error: 0	Decode Error: 0
Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0
SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0
SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0
Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0

Unclean Link (EP Side)

PCIe Demo for ISDS S/W

Help

Control PerformanceView PerformanceView Err Counter PerformanceView NoErr Counter PerformanceView SD Status LTSSM TraceView

Common Error counter(Max Value 65535)

Detect EI Infer:	0
Receiver Error:	0
Rx Recovery Request:	120
N_FTS Timeout:	0
Framing Error:	0
Deskew Error:	0
BAD TLP:	0
LCRC Error:	0
BAD DLLP:	0
Replay Num Rollover:	0
Replay Timeout:	0
Rx Nak DLLP:	0
Tx Nak DLLP:	0
Retry TLP:	0
FC Timeout:	1
Poisoned TLP:	0
ECRC Error:	0
Unsupported Request:	0
Completer Abort:	0
Completion Timeout:	0

Error counter per Lane(Max Value 15)

Lane0	Lane1	Lane2	Lane3
EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0
EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0
Decode Error: 0	Decode Error: 0	Decode Error: 0	Decode Error: 0
Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0
SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0
SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0
Rx Valid de-assertion: 2	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0
Lane4	Lane5	Lane6	Lane7
EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0	EBUF Over flow: 0
EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0	EBUF UnderRun: 0
Decode Error: 0	Decode Error: 0	Decode Error: 0	Decode Error: 0
Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0	Running Disparity Error: 0
SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0	SKP OS Parity Error: 0
SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0	SYNC Header Error: 0
Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0	Rx Valid de-assertion: 0

Clean 16GT/s Link Up

Clean Link-up



PCIe Demo for ISDS S/W

Help

Control PerformanceView PerformanceView Err Counter PerformanceView NoErr Counter PerformanceView SD Status LTSSM TraceView

No Error Counter Per Lane(Max Value 15)

Lane0	Lane1	Lane2	Lane3	Lane4	Lane5	Lane6	Lane7
EBUF SKP Add: 15	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0
EBUF SKIP DELL: 15	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0

LTSSM Counter(Max Value 4294967295)

L0 to Recovery Cnt: 1
L1 to Recovery Cnt: 0
Tx L0s: 0
Rx L0s: 0
ASPM L1 Reject: 0
L1: 0
L1 CPM: 0
L1.1: 0
L1.2: 0
L1 short duration: 0
L1.2 abort: 0
L2: 0
Speed Change: 1
Link width Change: 0

Rx Counter(Max Value 4294967295)

Rx Memory Write: 0	Rx Ack DLLP: 0
Rx Memory Read: 0	Rx Update FC DLLP: 12816994
Rx Config Write: 0	Rx Nullified TLP: 0
Rx Config Read: 0	Rx Duplicate TLP: 0
Rx IO Write: 0	
Rx IO Read: 0	
Rx Completion wo data: 0	
Rx Completion w data: 0	
Rx Message TLP: 0	
Rx Atomic: 0	
Rx TLP with Prefix: 0	

Tx Counter(Max Value 4294967295)

Tx Memory Write: 0	Tx Ack DLLP: 0
Tx Memory Read: 0	Tx Update FC DLLP: 12816650
Tx Config Write: 0	Tx Nullified TLP: 0
Tx Config Read: 0	
Tx IO Write: 0	
Tx IO Read: 0	
Tx Completion wo data: 0	
Tx Completion w data: 0	
Tx Message TLP: 0	
Tx Atomic: 0	
Tx TLP with Prefix: 0	

Clean Link-up (Time Elapsed)



PCIe Demo for ISDS S/W

Help

Control PerformanceView PerformanceView Err Counter PerformanceView NoErr Counter PerformanceView SD Status LTSSM TraceView

No Error Counter Per Lane(Max Value 15)

Lane0	Lane1	Lane2	Lane3	Lane4	Lane5	Lane6	Lane7
EBUF SKP Add: 15	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0	EBUF SKP Add: 0
EBUF SKIP DELL: 15	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0	EBUF SKIP DELL: 0

LTSSM Counter(Max Value 4294967295)

L0 to Recovery Cnt: 1

L1 to Recovery Cnt: 0

Tx L0s: 0

Rx L0s: 0

ASPM L1 Reject: 0

L1: 0

L1 CPM: 0

L1.1: 0

L1.2: 0

L1 short duration: 0

L1.2 abort: 0

L2: 0

Speed Change: 1

Link width Change: 0

Rx Counter(Max Value 4294967295)

Rx Memory Write: 0

Rx Memory Read: 0

Rx Config Write: 0

Rx Config Read: 0

Rx IO Write: 0

Rx IO Read: 0

Rx Completion wo data: 0

Rx Completion w data: 0

Rx Message TLP: 0

Rx Atomic: 0

Rx TLP with Prefix: 0

Rx Ack DLLP: 0

Rx Update FC DLLP: 15207636

Rx Nullified TLP: 0

Rx Duplicate TLP: 0

Tx Counter(Max Value 4294967295)

Tx Memory Write: 0

Tx Memory Read: 0

Tx Config Write: 0

Tx Config Read: 0

Tx IO Write: 0

Tx IO Read: 0

Tx Completion wo data: 0

Tx Completion w data: 0

Tx Message TLP: 0

Tx Atomic: 0

Tx TLP with Prefix: 0

Tx Ack DLLP: 0

Tx Update FC DLLP: 15207388

Tx Nullified TLP: 0

Wrap-up

Failure Manifestations



- **Link retrains on random interval**
 - Signal integrity
- **Link retrains on a regular interval**
 - Points to protocol problem – like missing SKPs, etc.
- **Other error indications**
 - Widely differing SKP add and delete across lanes
 - Multiple speed changes
 - Packet transmission counts not increasing
- **Frequency of errors**
 - Low frequency less concerning than high

- **Try forcing presets to eliminate link negotiation problems**
- **Force maximum link speed lower to see if problems remain at lower speeds**
- **Limit number of lanes (tape, interposer, etc.)**
- **Get a hardware analyzer in-system**

- **Look at details behind link retraining for clues to “mixed draft” links**
 - Complete failure to link at 16GT/s
 - Periodic retraining while at 16GT/s
- **Design to the latest draft**
 - 0.7 is solid
 - 0.9 as soon as it's published
- **Build in design-for-reliability / design-for-debug features to assist in link bring-up**

**Thank you for attending the
PCI-SIG Developers Conference Israel
2017.**

**For more information please go to
www.pcisig.com**